

Abstract: *This study was conducted on primary school students enrolled at primary schools under the administration of the federal Directorate of Education. The main purpose of this study was to analyze the extent to which intended critical thinking skill was developed in students. The objective of the study was to assess the student's critical thinking skills that had been developed through the curriculum. Male and female students from grade 3 to grade 5 were the population of the study. Data was collected from the students at primary schools administered by the Federal Directorate of Education. Primary-level students from the Islamabad sector comprised the study population. A multistage sampling technique was used to select samples for the study. Under descriptive research, this was cross-sectional research. Self-constructed criterion-referenced test was developed for students. Descriptive statistics were applied to analyze students' test data. The study concluded that there were significant differences in skill across grades 3, 4, and 5, with Grade 5 generally outperforming the others. This study emphasizes aligning assessments with curriculum outcomes, providing teachers with rubrics and professional development on skills-based assessments, and integrating technology to effectively assess 21st-century skills.*

Key Words: Assessment Skill, Critical Thinking Skill, Curriculum, Primary School Students, Education

Introduction

Critical thinking is the mental ability to learn things, make decisions, and offer unique solutions to problems. According to Paul (2017), it is the organized way of human thinking. Critical thinking has 2000 a 2000-year history when Socrates worked on learning through dialogue. Dewey called critical thinking as reflective ability in his book "How We Think". However, psychometric techniques are used globally for measuring critical thinking. Sarigoz (2012) investigated the critical thinking skills of students of primary and secondary level.

He concluded that students at the secondary level have better critical thinking skills than students at the primary level (Hu et al., 2016). According to Abrami et al. (2008), critical thinking is important for an individual's life. He argued that one cannot argue with asses and make decisions without critical thinking skills (Abrami et al., 2015). Edward Glaser is considered the father of critical thinking. Edward argued that critical thinking skills are essential skills in a progressive society as critical thinkers can logically think arguments and can reform societies. Literature on critical thinking skills is divided into three areas: education, psychology, and philosophy. These three areas have defined critical thinking from different perspectives. The psychological area stresses the expertise of skills and dispositions in multiple areas. It focuses on analysis, interpretation, logical thinking, and assessment skills (Gelder, 2005). Psychological areas focus on the process of critical thinking. Paul and Elder (2007) asserted that the philosophical area focuses on the personal qualities of individuals rather than outer behavior. It focuses on the rules of knowledge that are involved in the critical thinking process. Bailin and Siegel (2003) argued that critical thinking is not fixed and specific but varies according to subject areas. In the educational area, Benjamin Bloom has contributed a lot to critical thinking.

¹ Assistant Professor, Department of Education, Federal College of Education, Islamabad Pakistan.

² Assistant Professor, Faculty of Education, International Islamic University, Islamabad Pakistan.
 Email: Zarina.akhtar@iiu.edu.pk



Blooms offers a comprehensive hierarchy of cognitive thinking skills that educators excessively use. Higher-order cognitive skills like analysis, synthesis, and evaluation constitute critical thinking skills (Lai, 2011). There are varieties of definitions available for critical thinking. All definitions agree that critical thinking is something related to cognitive processes. However, there is disagreement on to which degree it is related to both cognition and dispositions (Ennis, 2011). Lai (2011) argued that the most mentioned critical thinking dispositions in literature are openness, flexibility, being up-to-date, fairness, and respect for others. American Philosophical Association addressed the definition and conceptualization of critical thinking skills. This Association prepared a Delphi report. Critical thinking is defined as Purposeful, self-regulatory judgment that leads to interpretation, analysis, evaluation, and inference, as well as explication of the evidentiary, conceptual, methodological, criteriological, or contextual considerations that underpin that judgment. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, and persistent in seeking results that are as precise as the subject and circumstances of inquiry allow (Facione, 2015).

Delphi report defined critical thinking as a purposeful activity that ultimately results in evaluation, interpretation, and self-control. Delphi report identified six cognition skills and two behavioral skills, which provide a framework for a comprehensive understanding of critical thinking skills. In 2012, the National Research Council reinforced the idea of the Delphi report. Paul and Elder (2007) argued that creativity and critical thinking are the two sides of the same coin. Both are dependent on each other. Both need each other for occurrence. Creativity is essential for becoming an effective critical thinker, and critical thinking enhances creative skills as well. Santos and Silva (2008) argued that there is no specific age for developing critical thinking. In a nutshell, the majority of researchers agreed that critical thinking involves two aspects, which are cognitive and dispositions. Assessing critical thinking skills is a complex process. However, the majority of researchers have recommended standardized tests and performance tests. Varieties of standardized tests are available for measuring critical thinking skills, such as the Watson-Glaser Critical Thinking Appraisal test and the California Critical Thinking Skills Test. Similarly, a variety of inventories are available to measure critical thinking dispositions.

Moreover, computer-based critical thinking tests are also available to measure students' grade-wise critical thinking skills. However, these tests are available mostly in multiple-choice question patterns. Ku (2009) argued that MCQ-based critical thinking tests only assess the students' recall and memorization skills, but the inner skills of students' reasoning cannot be explored. Performance-based tests include tasks in which students have to apply their knowledge and skills to new situations. Students have to do experiments, produce some product, or have to prepare some report. Rubrics are developed and students' performance is assessed according to that rubric. Many organizations have developed rubrics for performance-based tests (Becker, 2011).

The focus, Analyze, Infer, and Decide (FAID) model of critical thinking, developed by Paul and Elder (2006), is a comprehensive framework that identifies four core skills essential for critical thinking. These skills are focusing questions (the ability to identify and clarify the central issue or problem, asking relevant and insightful questions to frame the inquiry), analysis (the ability to break down complex information into component parts, identifying patterns, relationships, and causes, as well as evaluating evidence and arguments), inference (the ability to draw logical conclusions based on the analysis, making sound judgments and recognizing implications) and deciding on action (the ability to select a course of action based on the conclusions drawn, considering multiple perspectives, evaluating consequences, and making a decision that is well-reasoned and ethical). However, various challenges are associated with the measurement of critical thinking skills. Wilson et al. (2012) argued that there are various steps involved in constructing a critical thinking test.

These steps include defining the test construct, selecting test items, determining the test scores, considering the test administration issues, selecting the targeted sample, administering the test, analyzing scores, and preparing test result reports. However, Ku (2009) argued that defining the construct of the test is a very complex task and needs the special attention of the test developer. Care et al. (2018) argued that there are certain things that should be considered at the time of constructing a critical thinking test. These

include task authenticity, ill-defined tasks, open-ended questions, using prompts, and measurable test objectives.

Methodology

By method, it is descriptive research. Descriptive research focuses on what questions rather than why questions. Descriptive research is a cross-sectional type of research. Cross-sectional research focuses on population data across different sections at the same time. All students from grade 3 to grade 5 were the population of study. Data was collected from primary schools under the administration of the Federal Directorate of Education. Students from the Islamabad sector comprised the study population. There are a total of 189 primary schools administered by the Federal Directorate of Education. Among them, 90 schools are working in Islamabad Area. The total number of students (grade 3 to grade 5) in 90 schools was 7000. Therefore, the population size was 7000 students. The researcher used the Cluster Sampling technique. The cluster sampling technique is a type of probability sampling. The researcher divides the whole population into clusters/groups. The researcher used multi-stage sampling to reach the clusters for selecting samples. A multistage sampling technique was used to select a sample for the study. In multi-stage sampling, the whole population is divided into small groups to get accurate and fair data. It makes the larger population manageable (Lohr, 2021). A sample size of 10% was selected from the targeted population. Therefore, a simple size of approximately 700 was selected for data collection. In the first stage, schools were selected for data collection.

Schools of Islamabad (sector G and I) were selected for data collection. Islamabad, the capital city of Pakistan, is divided into eight zones, each further subdivided into sectors. The Islamabad Sector system is a well-planned and organized way of categorizing the city into manageable parts. Each sector is identified by a letter (B, C, D, E, F, G, H, I.) and is designed to provide a unique identity to each area. This system helps in efficient urban planning, administration, and navigation. The sectors are typically self-contained, with their own commercial and residential areas, parks, mosques, and other amenities. 3 schools from Sector G and three schools from Sector I were selected. In the second stage, students were selected from these schools. Students in grades 4 and 5 were considered for data collection.

Instrument

Self-constructed criterion-referenced test was developed for students. A criterion-referenced test was constructed to measure each skill. For critical thinking skills, mathematics and English language subjects were considered. A one-way table of specifications was developed to develop students' tests for measuring critical thinking skills. A pilot test of all students' tests was carried out. A one-way table of specifications was developed for constructing the test. Item analysis was carried out for a test. The reliability and validity of the test were also determined by the researcher. Reliability of the test is defined as the extent to which the results of the test are considered stable and consistent. The reliability of students' tests was estimated by calculating the Pearson coefficient. Criterion and content validity were determined for students' tests. The researcher also got expert opinions for determining test validity. Subject specialists were requested to review the drafts of the test to ensure the content validity of the test.

Data Collection

Data were collected from primary-level students (grade 3 to grade 5). The researcher first got official permission from relevant authorities. School principals were informed about the purpose of the research and requested data collection. The purpose of the research was shared with students, concerned teachers, and school principals. After seeking their permission, self-constructed tests were administered once to the students from grade 3 to grade 5. Data collection took about half a month. Paper pencil tests were administered once with the selected sample. After getting their consent, the researcher collected data.

Data Analyses

Mean score analysis and ANOVA analyses were used to analyze the data.

**Table 1**

Comparison of mean score on Critical Thinking Skill

Classes	N	Mean	Standard deviation
3	228	11.18	1.56
4	228	11.06	1.64
5	228	11.27	1.45

Table 1 provides a comprehensive insight into the distribution characteristics of a variable across three distinct classes. Each class comprises 228 observations, ensuring a robust comparison. Upon analysis, it becomes evident that there are subtle differences in the mean scores across the classes. Class 5 emerges with the highest mean score of 11.27, indicating that, on average, this class exhibits the highest values of skills development among the three. Conversely, Class 4 demonstrates the lowest mean score at 11.06, suggesting a slightly lower average value compared to the other classes. Notably, Class 3 falls between Classes 4 and 5 in terms of mean score, with a value of 11.18. However, the standard deviation within each class adds nuance to the interpretation. Class 5 exhibits the lowest standard deviation of 1.45, indicating a relatively tighter spread of data points around the mean compared to the other classes. Conversely, Class 4 displays the highest standard deviation of 1.64, signifying a greater variability in the data within that class. This implies that while Class 5 has the highest mean score, its data points are more tightly clustered around the mean, potentially indicating a more consistent performance or outcome within that class. In contrast, Class 4, despite having a lower mean score, shows a wider dispersion of data points, suggesting a greater variability in the observed variable.

ANOVA analyses on Critical Thinking skill

There is no difference between critical thinking and a national curriculum

Table 2

Groups	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	31.30	2	15.65	6.56	.001
Within Groups	1620.28	680	2.38		
Total	1651.58	682			

The summary of the results in the above table, containing $F=6.568$ and $p=.001$, indicates that there exists a statistically significant relationship between critical thinking skills and national curriculum. Hence, the null hypothesis was accepted. The value of the F ratio (6.56) and respected P value (.001) indicates that there is a significant difference among the students of different classes. Hence, the value of the F ratio (6.56) and respected p-value (.001) indicates that there is a significant difference among the students of different classes.

Table 3

Post hoc Table on Critical Thinking Skill

Comparison among classes	Mean Difference	Sig.
Grade 3 V.S Grade 4	.14564	.315
Grade 3 V.S grade 5	.36313*	.012
Grade 4 V.S Grade 5	.50877	.000

Post hoc test statistics provided a significant mean score difference between the groups, whereas the difference in critical thinking skills mean score between grade 3 and grade 4 was not significant. Similarly, the difference in critical thinking skills mean score between grade 4 and grade 5 was not significant. However, there was a significant mean score difference between the critical thinking skills of grade 3 and grade 5.

Findings, Discussion, Conclusions, and Recommendations

The mean scores for critical thinking skills across different grade levels indicate minimal variations. Grade 5 students have the highest mean score ($M = 11.27$, $SD = 1.45$), followed by Grade 3 ($M = 11.18$, $SD = 1.56$), with Grade 4 showing the lowest mean score ($M = 11.06$, $SD = 1.64$). Therefore, according to these mean scores, Grade 5 students appear to have slightly better critical thinking skills compared to students in Grade 3 and Grade 4. The analysis of variance (ANOVA) results demonstrates significant differences among the critical thinking skills of students of 3 classes ($F(2, 68) = 6.56$, $p = .001$). Notably, the between-groups variation amounted to 31.301 units, with a mean square value of 15.65, while within-groups variation totaled 1620.28 units. This discrepancy suggests that the variability observed between the groups is statistically significant compared to the variability within the groups. These findings indicate notable distinctions among the groups under examination. The post hoc analysis of the critical thinking skills of students' of 3 classes reveals interesting distinctions among grade levels. When comparing Grade 3 to Grade 4, the mean difference was found to be .14564, with a non-significant p-value of .315, suggesting no statistically significant difference between these groups. However, a significant contrast emerged between Grade 3 and Grade 5, with a notable mean difference of .36 ($p = .012$), indicating a substantial disparity in performance between these grades. Similarly, a significant difference was observed between Grade 4 and Grade 5, where the mean difference was .50, with a p-value of .000. These findings underscore significant variations in skill proficiency across different grade levels, particularly emphasizing the substantial gap between Grade 3 and Grade 5.

The result analysis revealed significant differences in critical thinking skills among students in grades 3, 4, and 5, indicating that the curriculum has a significant impact on the development of this skill. Moreover, results show that there is a significant mean score difference between grade 3 and grade 5, and also between grade 4 and grade 5, highlighting the importance of curriculum in developing critical thinking skills. A study by Abrami et al. (2015) found that students who received explicit instruction in critical thinking skills showed significant improvement in their critical thinking skills and problem-solving skills. Another study by Halpern (2014) found that students who were taught critical thinking skills as part of a broader critical thinking program showed improved critical thinking skills and problem-solving skills. It was concluded that Grade 5 students possess slightly better critical thinking skills compared to Grade 3 and Grade 4 students. It was concluded that significant differences exist in critical thinking skills among students in three different classes (Grade 3; Grade 4, and Grade 5). The results showed that the variability between the groups was statistically significant, indicating that the groups have distinct levels of critical thinking skills.

Recommendations

1. Grade 4 students have the lowest score in critical thinking skills. Interventions focusing on critical thinking skills development may benefit Grade 4 students.
2. Teachers may be provided with rubrics and scoring guides for student learning as it will help them in effective evaluation of students' performance in the skill. Studies may be conducted on exploring multiple assessment methods for measuring 21st century skills. These methods can be both qualitative and quantitative.
3. As this study was descriptive in nature, longitudinal studies may be conducted on assessing 21st century skills across different time frames.

References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*, 85(2), 275–314. <https://doi.org/10.3102/0034654314551063>
- Bailin, S., & Siegel, H. (2003). *Critical thinking*. USA: Blackwell Publishing.
- Becker, A. (2011). Examining rubrics used to measure writing performance in U.S. intensive English Programs. *The CATESOL Journal*, 22(1), 113–130. <https://doi.org/10.5070/b5.36219>
- Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice*, 32(3), 179–186. <https://doi.org/10.1080/00405849309543594>



- Facione, P. A. (2015). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. American Philosophical Association.
- Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College Teaching*, 53(1), 41–46. <https://doi.org/10.3200/CTCH.53.1.41-48>
- Hu, W., Jia, X., Plucker, J. A., & Shan, X. (2016). Effects of a critical thinking skills program on the learning motivation of primary school students. *Roeper Review*, 38(2), 70–83. <https://doi.org/10.1080/02783193.2016.1150374>
- Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4(1), 70–76. <https://doi.org/10.1016/j.tsc.2009.02.001>
- Lai, E. R., & Viering, M. (2012). *Assessing 21st Century Skills: Integrating Research Findings*. Pearson.
- Lohr, S. L. (2021). *Sampling: design and analysis*. Chapman and Hall/CRC.
- Paul, R., & Elder, L. (2006). Critical thinking: The nature of critical and creative thought. *Journal of Developmental Education*, 30(2), 34–36. <https://files.eric.ed.gov/fulltext/EJ986272.pdf>
- Paul, R., & Elder, L. (2007). *A guide for educators to critical thinking competency standards: Standards, principles, performance indicators, and outcomes*. Foundation for Critical Thinking Press.
- Santos, D., & Silva, G. V. (2008). Authenticating materials through critical thinking: the case of teaching and learning suggestions in Portuguese. *Hispania*, 110–123. <https://www.jstor.org/stable/20063627>
- Sarigoz, O. (2012). Assessment of the high school students' critical thinking skills. *Procedia, Social and Behavioral Sciences*, 46, 5315–5319. <https://doi.org/10.1016/j.sbspro.2012.06.430>
- Wilson, M., Bejar, I., Scalise, K., Templin, J., Wiliam, D., Irribarra, D.T. (2012). Perspectives on Methodological Issues. In: Griffin, P., McGaw, B., Care, E. (eds) *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2324-5_3